融合改进UNet和迁移学习的棉花根系图像分割方法

唐 辉[1], 王 铭[2], 于秋实[1], 张佳茜[1], 刘连涛[3], 王 楠[1*]

(1. 河北农业大学 机电工程学院,河北保定071001,中国; 2. 河北省教育考试院,河北石家庄050091,中国;
3. 河北农业大学 农学院,河北保定071001,中国)

摘 要: [**目的/意义**] 根系是植物组成的重要部分，其生长发育至关重要。根系图像分割是根系表型分析的重要方法，受限于图像质量、复杂土壤环境、低效传统方法，根系图像分割存在一定挑战。[**方法**] 为提高根系图像分割的准确性和鲁棒性，本研究以UNet模型为基础，提出了一种多尺度特征提取根系分割算法，并结合数据增强和迁移学习进一步提高改进UNet模型的泛化性和通用性。首先，获取棉花根系单一数据集和开源多作物混合数据集，基于单一数据集的消融试验测试多尺度特征提取模块（Conv_2+Add）的有效性，与UNet、PSPNet、SegNet、DeeplabV3Plus算法对比验证其优势。基于混合数据集验证改进算法（UNet+Conv_2+Add）在迁移学习的优势。[**结果和讨论**] UNet+Conv_2+Add相比其他算法（UNet、PSPNet、SegNet、DeeplabV3Plus），mIoU、mRecall和根系F_1调和平均值分别为81.62%、86.90%和78.39%。UNet+Conv_2+Add算法的迁移学习相比于普通训练在根系的交并比（Intersection over Union，IoU）值提升1.25%，根系的Recall值提升1.79%，F_1调和平均值提升0.92%，且模型的整体收敛速度快。[**结论**] 本研究采用的多尺度特征提取策略能准确、高效地分割根系，为作物根系表型研究提供重要的研究基础。

关键词: 深度学习；根系图像分割；UNet；多尺度特征；迁移学习

中图分类号: S24;TP181 **文献标志码:** A **文章编号:** SA202308003

引用格式:唐辉, 王铭, 于秋实, 张佳茜, 刘连涛, 王楠. 融合改进UNet和迁移学习的棉花根系图像分割方法[J]. 智慧农业(中英文), 2023, 5(3): 96-109. DOI：10.12133/j.smartag.SA202308003

TANG Hui, WANG Ming, YU Qiushi, ZHANG Jiaxi, LIU Liantao, WANG Nan. Root image segmentation method based on improved UNet and transfer learning[J]. Smart Agriculture, 2023, 5(3): 96-109. DOI：10.12133/j. smartag. SA202308003 (in Chinese with English abstract)

1 引 言

根系作为植物和外界环境交换的器官，包括代谢、吸收、矿物和有机物交换等，地上部植株的生长也受到地下根系影响[1,2]。植株品种培育改良，对提高作物产量和养分资源利用率等方面具有重要意义，也是第二次绿色革命的重要内容[3]。由于土壤的遮蔽性和根系复杂生长状态，根系表型获取仍是当前研究热点。

传统的根系表型获取依靠人工挖掘[4]，耗费较高的物资人力，挖掘过程耗时很长，且获取的根系上会造成细小的根段缺失，无法满足当前的根系表型研究。随着表型技术发展，根系表型获取转向依靠视觉穿透效果的生长介质获取根系图像，主要方法有气培法、水培法、凝胶培养法和纸基培养法等[5]，但是透明介质方法根系观测结果与实际土壤观测结果差距大。随着高分辨率成像设备的普及使

收稿日期：2023-07-28

基金项目：河北省教育厅青年拔尖人才计划项目（BJ2021058）；中央引导地方科技发展资金项目（236Z7402G）；华北作物改良与调控国家重点实验室自主课题（NCCIR2021ZZ-23）

作者简介：唐 辉，研究方向为深度学习、根系表型。E-mail：318901249@qq.com

*通信作者：王 楠，博士，副教授，研究方向为表型组学、图像处理。E-mail：cmwn@163.com

用和成本的下降，基于成像装置土壤根系表型获取方法能够较好地还原根系生长状态、采集速度更快、成像分辨率更高。较早的成像装备观测方法起源“微根管”[6]技术，将微根管装置深埋地下观测根系结构，该技术已在多种作物上得到应用。但是装置的放置位置会干扰根系生长，并且只能采集到局部根系图像。随着探测技术在医学领域中应用，一些学者将医学设备应用到根系表型研究中[7]，如穿透射线成像法X射线计算机断层扫描（Xray Computed Tomography，XCT）和核磁共振成像（Magnetic Resonance Imaging，MRI）[8,9]。但这些方法无法精准识别细根，且受到土壤含水量的干扰。射线设备和微根管设备使用成本大、维护成本高、成像分辨率普遍低。近年来数码设备成像法得到了更多应用，数码装置相对于其他成像设备成本相对低，成像分辨率可以调节，得到根系图像更精确。Hammac等[10]使用数码设备实现油菜、亚麻和扁豆两个水分水平下根毛发育特征。Mohamed等[11]采用了扫描设备完成了对核桃树根的扫描，且实验组前期设计了RhizoPot[12]装置能够更高效获取根系图像。获取原位根系图像后，需要提取根系表型特征。传统方法的根系识别包括人工描绘、半自动交互识别和全自动阈值分割。人工描绘存在识别效率低、工作量大、结果误差高等问题。半自动交互识别是研究者依据视觉观察，通过辅助软件进行图像识别。虽然半自动交互式可达到较高精度，但过于依赖观测者分辨根系的主观能力和自身经验，单张复杂根系图像分割需4～5 h，效率过低，难以实现高通量原位根系图像分析。基于全自动阈值分割的图像处理方法提高了根系识别效率，如Digital Imaging of Root Traits（DIRT）[13]，General Image Analysis of Roots （GiA Roots）[14]，IJ-Rhizo[15]和EZ-Rhizo[16]等可以自动识别根系，但容易受到土壤噪声影响，识别准确率低。基于深度学习的语义分割网络给根系识别提供了更好的方法。基于语义分割的根系识别，通过反复前向迭代与反向传播过程不断调整网络各层权重，提取多尺度像素特征，实现土壤背景与根像素的二值化分离。一般的卷积神经网络虽可区分根系与土壤，但网络局限于单一尺度预测，无法处理根系尺寸变化问题，导致图像边缘细节缺失。2015年全卷积神经网络（Fully Convolutional Networks，FCN）的提出实现像素级分类[17]。FCN在卷积神经网络的基础上，增加了上采样和跳跃结构实现图像的像素级分割并提高分割鲁棒性。Kamal等[18]基于该模型实现了杂草和作物的分割。但FCN存在处理结果不够精细和没有考虑像素与像素间关系等问题。SegNet[19]在FCN的基础上采用对称的编码器-解码器结构，添加了非线性上采样，实现端到端训练，训练结果优于FCN。Wang等[20]基于SegNet模型设计了高通量根系分析软件SegRoot，实现了根系与土壤的区分。PSPNet是基于金字塔池化模块实现的[21]，聚合多感受野的上下文信息，提高了获取全局信息的能力。Zhang等[22]基于改进PSPNet实现农业地区分割。谷歌公司提出的DeepLabV3plus模型[23]，在V3[24]基础上引入了编码器和解码器结构，并使用深度可分离卷积减少模型参数量、提高准确率。本研究在前期研究中基于V3plus网络，通过改进上采样方式实现了根系有效处理[25,26]。UNet模型的提出用来解决医学领域分割问题[27]。根系类似人体血管组织，本研究通过改进UNet模型来实现根系的分割。由于其出色的分割性能，该模型也被广泛应用到其他领域，如城市植被提取[28]、地表覆盖物分类[29]等。迁移学习是利用已训练的模型投入到其他相关领域中重新训练模型参数的方法。基于该方法能够减少模型训练的时间，增加模型泛化能力[30]。本研究通过改进UNet模型来实现根系识别，基于单一数据集和混合数据集的数据增强和迁移学习方法实现根系的高效分割，提高改进模型泛化能力，为根系表型分析提供新的方法。

2　试验材料

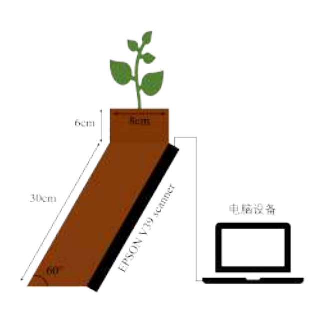
本模型采用两种数据集进行试验。单一数据集是基于数码扫描设备采集的棉花原位根系图像，混合数据集为网络公共开源的多作物微根管根系数据集。

2.1 单一数据集

单一数据集获取位于中国河北省保定市河北农业大学实验站（38.85°N，115.30°E）。根系采集对

象是棉花，品种为国欣棉9号、转基因抗虫棉K836。棉花作物通过根系获取装置（RhizoPot）进行种植，该装置使用透明亚克力板围成梯形透明根室，两侧倾斜角与地面呈67.5º，并安装数码成像扫描仪（Epson PerceptionV39，Suwa，Japan），除梯形两侧外，周围用黑色遮蔽纸围住防止阳光对根系产生干扰。RhizoPot装置竖直高度为350 mm、上边长490 mm、下边长265 mm、宽度205 mm，容积共14.5 L。扫描仪装置通过Software Development Kit （SDK）（Epson Scan SDK V10102_20210510）控制连接电脑，一台电脑可同时控制多台RhizoPot装置。图像获取装置如图1所示。根系图像从第1天到第110天内进行连续拍摄，图像采集像素大小为1200 dpi，分辨率像素为10,200×14,039，图像的深度为24位。棉花根系图像共110张，排除掉图像中包含噪声和图像不清晰的，剩余100张。共拍摄六组照片，每组100张，共600张。

建立数据集，在获取的图像中随机选择100张进行标注。图像标注使用的是Adobe Photoshop CC2020 （Adobe Inc.，San Jose，CA，United States）。首先使用Adobe Photoshop打开图像，并建立新图层，使用套索工具选择要标注的根系，然后使用油漆桶工具为选择的根系填充白色。重复上述操作直到所有根系标注完成，再将土壤背景填充为黑色，最后保存图像为8位的*.png图像。图像标注示例如图2所示。每张图像的标注时间约3 h。由于训练集图片的分辨率过大，本研究采用滑窗方法分割为分辨率像素为768×768的小图训练，不足的部分使用黑色填充。分割后的数据集按照7∶2∶1的比例划分为训练集、验证集和测试集，四舍五入划分后数量分别为19,360、5531和2766张。



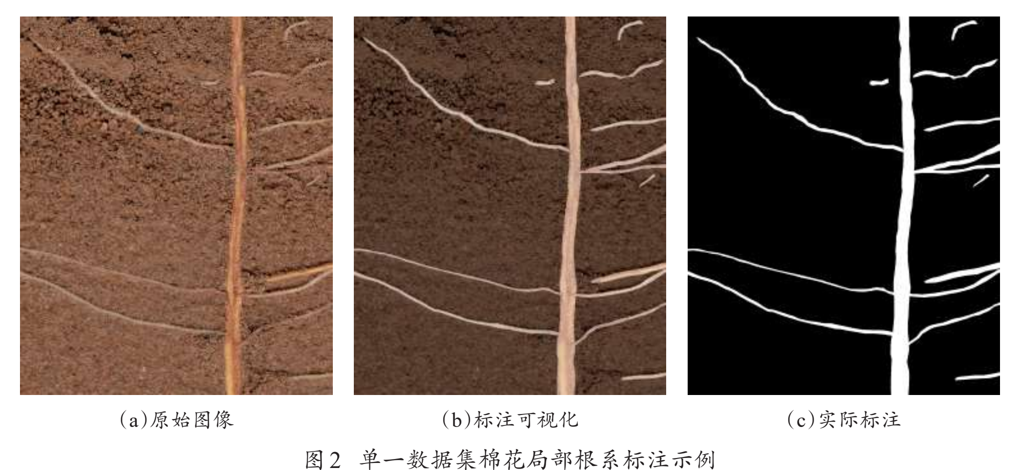
(a)RhizoPot



(b)多扫描仪采集

图1 单一数据集棉花根系图像获取装置

Fig. 1 Single dataset cotton root system image acquisition device



(a)原始图像　(b)标注可视化　(c)实际标注

图2 单一数据集棉花局部根系标注示例

Fig. 2 Example of localized root annotation for cotton in a single dataset

2.2 混合数据集

混合数据集为公共开源的微根管数据集，该数据集来源于PRMI论文[31]，数据集网址为https://gatorsense.github.io/PRMI/。原始的数据集包含不同时间段和不同土壤深度的棉花、木瓜、花生、芝麻、

向日葵5种作物图像，且该数据被划分为训练集和测试集2类。为了适用于本试验，将原始的训练集重新划分为训练集、验证集。三种数据集详细信息如表1所示。

表1　开源混合数据集的作物种类、像素大小及包含的张数

Table 1　Crop types, pixel sizes, and number of sheets in open source mixed dataset

数据集类型	棉花	木瓜	花生	花生	芝麻	芝麻	向日葵
分辨率/px	736×552	736×552	640×480	736×552	640×480	736×552	640×480
训练集/张	1271	282	10,087	11,485	1438	8637	2211
验证集/张	564	131	3413	3347	318	2625	722
测试集/张	577	133	3542	4793	404	3048	967

2.3 数据增强

由于作物生长在不同的环境中，采集的根系图像也存在差异，如土壤颜色差异、根系的粗细、根系的生长角度，所以试验采用的数据增强方法可以实现样本的多样性。

试验采用四种图像处理方法：

（1）改变图像属性。随机改变图像的亮度，幅度为0.5～1.5倍；随机变化图像的对比度，幅度为0.5～1.5倍；随机变化图像的饱和度，幅度为0.5～1.5倍。

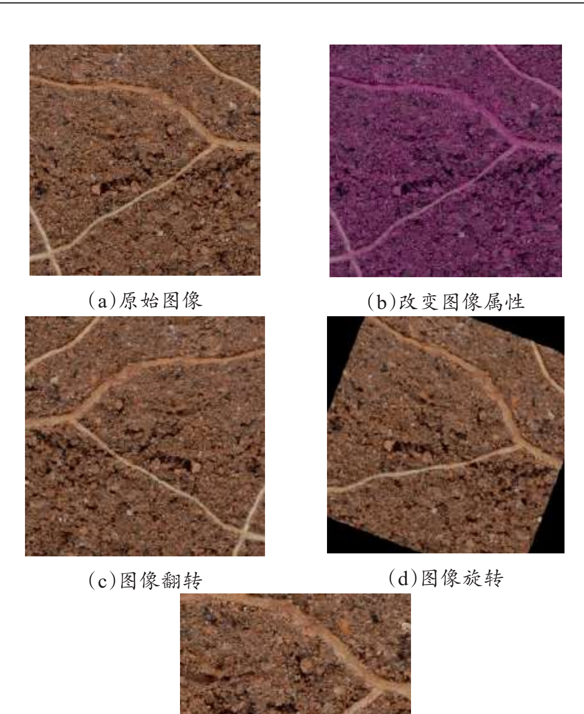
（2）图像随机裁剪。随机裁剪原始图像的某一区域，裁剪的范围为随机高或宽的0.2～1倍，最后将裁剪后的图像放大，分辨率像素为512×512。

（3）图像旋转。随机对图像进行−90°～90°的图像旋转。

（4）图像翻转。随机对图像按照x轴或y轴进行翻转。

在模型训练时随机读取训练数据集内图像，读取后对图像按照设定几率开展数据增强操作。四种数据增强方法几率设置为0.3，且每个方法都互不影响。读取图像时程序会设置随机数，当随机数小于0.3时图像才会进行变化。使用该方法在模型每代训练中图像都存在差异，使得样本多样性进一步增大。

四种图像增强的方法示例如图3所示。



(a)原始图像　(b)改变图像属性　(c)图像翻转　(d)图像旋转　(e)图像随机裁剪

图3　模型训练的原始图像及4种不同的数据增强方法结果

Fig. 3　Original image of model training and the results of four different data enhancement methods

3　研究与方法

3.1 模型结构

UNet与其他的卷积模型相似，采用U型编码-解码器结构，编码器主要通过连续的卷积和池化进行特征提取，解码器用于对特征图的信息还原，恢复特征图到原尺寸大小。改进后的模型主要在编码器和解码器之间的跳跃连接添加了额外的提取模块。

UNet编码器包含五层，前四层是由卷积-最大池化进行特征提取并下采样，然后保留每层的特征图用于特征融合，编码器最后一层只包含卷积层。编码器每层的特征图提取计算方法如公式（1）所示。

$$\bar{C}_i = MP_{i_2}\left(2 \times \left(Conv\left(C_{i_1}\right)\right)\right)$$
$$i_1 \in [1,5] \, and \, i_2 \in [1,4] \quad (1)$$

其中，*Conv* 表示一次完整卷积运算，包含卷积、*Batch Normalization* 归一化和 *ReLU* 激活函数；*MP* 表示最大池化层，且只作用在1～4层；C_{i_1} 表示每层的输入特征图；$\bar{C}_i$ 表示第 *i* 层的输出特征图。

原始的UNet模型在结构上利用下采样完成不同图像尺寸的卷积特征提取，但在池化操作时难免会造成信息损失。为减少信息的损失，本研究提出的多尺度图像特征提取融合模块完成细节信息弥补。该模块主要放置于整体模型的中间三层。首先对原始的输出图像进行双线性插值得到三种尺寸不同的图像（256×256 dpi，128×128 dpi，64×64 dpi），然后通过卷积提取特征图并将其与相对应跳跃连接层的特征图相加，然后与上采样的特征图拼接完成对细节信息的弥补。多尺度特征提取的计算方法如公式（2）所示。

$$\hat{C}_i = Add\left(2 \times \left(Conv\left(\varnothing_i(C)\right)\right), \bar{C}_i\right) \quad i \in [2,4] \quad (2)$$

其中，*C* 表示原始的图像；$\varnothing_i$ 表示对原始图像进行不同尺度双线性插值操作；*Add* 表示对两种特征图进行加操作；$\hat{C}_i$ 表示模块输出特征图。

编码器层和多尺度图像特征提取融合模块计算完成后，由解码器完成融合还原，解码器每层特征图融合计算见公式（3）。

$$\bar{C}_o = 2 \times \left(Conv\left(Concat\left(UP\left(\bar{C}_{i-1}\right), \hat{C}_i\right)\right)\right) \quad i \in [1,4] \quad (3)$$

其中，*UP* 表示上采样函数；*Concat* 用于拼接特征图；$\bar{C}_{i-1}$ 表示上一层输出特征图；$\bar{C}_o$ 表示本层的输出特征图。

模型改进后的整体结构如图4所示。

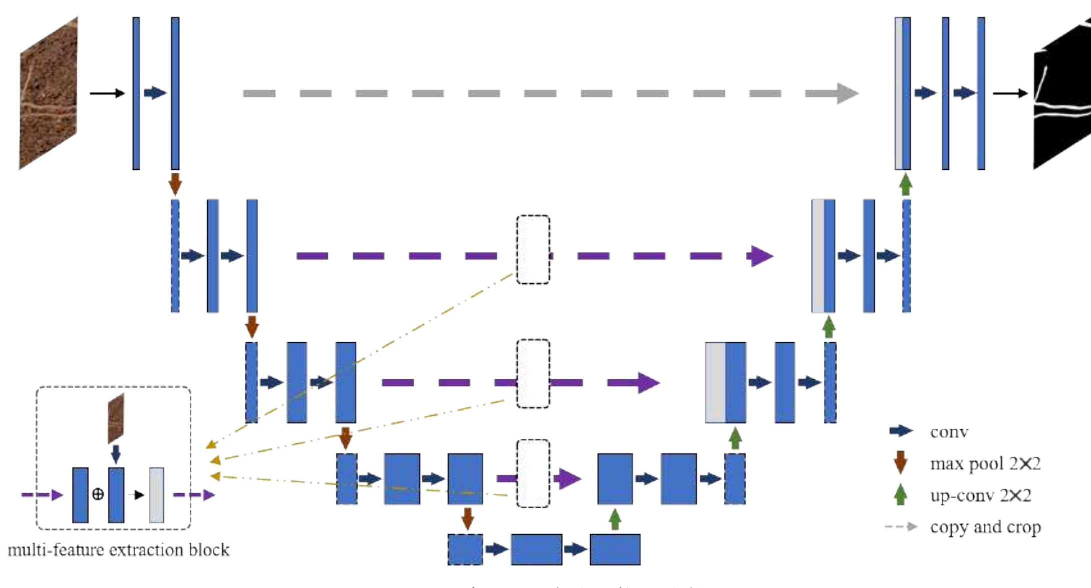


图4 改进UNet模型的整体结构

Fig. 4 Overall structure of the improved UNet model

3.2 研究策略

试验基于单一数据集和混合数据集完成对改进UNet模型的训练和预测，采用数据增强、模型改进和迁移学习3种方法提高了根系的分割能力和模型的泛化能力。

试验的主要过程基于单一数据集选择改进最佳模型，并与其他算法对比验证其优势，使用迁移学习和混合数据集进一步提高模型泛化能力，上述所有模型训练过程中全都采用数据增强的方法，其技术路线如图5所示。

为实现根系图像分割算法训练、测试和对比，本研究对三种试验提出以下策略。

（1）消融实验策略。卷积和特征融合对多尺度信息提取至关重要，为改善原始模型在编码器层下采样的信息损失，本研究对原始模型设计了几种不同的多尺度图像特征提取融合方式，其结构主要包括图像的特征提取方式和特征融合方式两方面。在

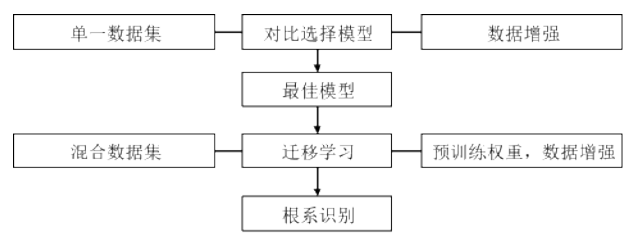


图5 根系分割算法研究技术路线

Fig. 5 Technical route for the root segmentation algorithm

特征提取上改进包含：注意力机制（Convolutional Block Attention Module，CBAM）、深度分离卷积（DP Conv）和普通卷积（Conv），在特征融合方式上分为拼接（Concat）和相加（Add）。主要的对比策略如表2所示，Conv_2+Add是最佳策略。消融实验是基于单一数据集，训练过程使用数据增强，模型采用随机权重加载。

表2 消融实验的模型改进策略及解释

Table 2 Model improvement strategies and explanations for ablation experiments

模型改进策略	模型解释
Conv_1+Concat	进行一次完整卷积计算,并将其上采样和每层编码器中特征图拼接
Conv_2+Concat	进行两次完整卷积计算,并将其上采样和每层编码器中特征图拼接
DP Conv+Concat	进行两次深度可分离卷积计算,并将其上采样和每层编码器中特征图拼接
CBAM+Concat	进行完整卷积后再进行注意力机制计算,并将其上采样和每层编码器中特征图拼接
Conv_1+Add	进行一次完整卷积计算,并将其与跳跃连接的特征图相加并和上采样进行拼接
Conv_2+Add（本研究）	**进行两次完整卷积计算,并将其与跳跃连接的特征图相加并和上采样进行拼接**
DP Conv+Add	进行两次深度可分离卷积计算,并将其与跳跃连接的特征图相加并和上采样进行拼接
CBAM+Add	进行完整卷积后,进行注意力机制计算,并将其与跳跃连接的特征图相加并和上采样进行拼接

（2）对比试验策略。对比试验验证改进最佳算法优越性，对比模型有PSPNet、SegNet、DeeplabV3Plus，模型基于单一数据集训练，训练过程使用数据增强，模型采用随机权重加载，未使用迁移学习。

（3）迁移学习策略。迁移学习是进一步提高模型泛化能力和根系分割能力，训练模型有原始UNet和改进最佳算法，基于混合数据集训练，训练过程使用数据增强，模型采用单一数据集预训练权重加载。

3.3 研究条件

上述所有模型的训练及推理都基于Ubuntu22.04系统完成，系统的配置的处理器为intel i5-12400F（2.5 GHz）和32GRAM，显卡型号为GTX 3080Ti，显存为12 GB。训练使用的深度学习框架是Pytorch11.6。

在训练时为保证公平性，所有模型的训练采用相同的训练环境，模型的超参数设置也相同。训练全部采用Adaptive Moment Estimation with decoupled weight decay（ADAMW）优化器，参数betas分别为0.9和0.999，初始的学习率设置为0.0001，学习率衰减策略采用余弦退火学习率衰减策略（Cosine Annealing LR），模型训练共100代。

3.4 模型的评估

为了验证模型性能，需要借助评估指标对模型进行评估。在本研究中根系和土壤背景相当于对图像的逐像素分类，因此需要借助混淆矩阵来统计分类结果和实际值，从而进一步获取多种评估指标。本研究采用4种评估指标为准确率（Precision）、召回率（Recall）、交并比（Intersection over Union）和F_1，计算方法如公式（4）～（7）所示。

$$Precision = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{4}$$

$$Recall = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{5}$$

$$IoU = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN} + \mathrm{FP}} \tag{6}$$

$$F_1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{7}$$

其中，TP表示正确预测为根的根像素数；FP表示被预测为根的背景像素数；FN表示被预测为背景的根像素数；TN表示正确预测为背景的背景

像素数。IoU可以评估像素分类结果与实际值之间的相似性。*Precision*、*Recall*用于验证像素分类的正确率，F_1则是*Precision*、*Recall*的调和平均值。四种评估标准的取值范围都在0～1之间。

4 结果与分析

4.1 消融实验

为验证改进模型中各部分对模型的性能影响，对卷积方式和融合方式进行消融实验，获取各种改进模型评估指标值。每个模型都使用相同的试验参数进行训练，训练结束后使用测试集对模型进行评估，获取各项评估指标。消融实验评估指标结果如表3所示。对比原始模型发现，Conv_2+Add综合评估指标效果最佳，平均交并比（mean Intersection over Union，mIoU）、平均召回率（mean Recall，mRecall）和根系F_1值分别为81.62%、86.90%和78.38%，平均准确率（mean Precision，mPrecision）值为91.12%相比于原始模型有所降低。但鉴于Precision和Recall是一对相对矛盾的指标，根据F_1调和平均值的结果改进后模型还是相比于原始模型有所提升，而且根系的IoU值相比于其他改进模型最高。

表3 消融实验各种改进模型的评估指标

Table 3 Evaluation Indicators for various improved models in ablation experiments

评估指标	UNet	Conv_1+Concat	Conv_2+Concat	DP Conv+Concat	CBAM+Concat	Conv_1+Add	Conv_2+Add	DP Conv+Add	CBAM+Add
R IoU/%	63.71	55.61	63.83	62.81	63.71	63.39	64.44	63.00	63.90
B IoU/%	98.79	98.56	98.79	98.76	98.79	98.77	98.79	98.75	98.79
mIoU/%	81.25	77.08	81.31	80.79	81.25	81.08	81.62	80.88	81.35
R Recall/%	72.39	61.68	72.68	71.18	72.22	72.57	74.25	72.36	72.89
B Recall/%	99.60	99.68	99.59	99.61	99.60	99.57	99.55	99.56	99.56
mRecall/%	85.99	80.68	86.13	85.39	85.91	86.07	86.90	85.96	86.24
R Precision/%	84.16	84.95	83.98	84.23	84.39	83.36	83.00	82.96	83.82
B Precision/%	99.18	98.87	99.19	99.15	99.18	99.19	99.24	99.18	99.20
mPrecision/%	91.67	91.91	91.59	91.69	91.78	91.27	91.12	91.07	91.51
R F_1/%	77.83	71.47	77.92	77.16	77.83	77.59	78.38	77.30	77.97
B F_1/%	99.39	99.27	99.39	99.38	99.39	99.38	99.39	99.37	99.38

注：R代表根系；B代表背景

由图6可以看出，模型全部使用随机梯度下降的优化算法，由于其高效的计算能力，使得各模型在训练时损失达到收敛的时间差距不大（DP Conv+Add除外），都在训练到10代之后模型损失趋于平稳。

在对比Concat和Add两种方式上的图像特征提取发现，使用双层卷积得到的指标最好，相比于单层卷积方法，双层卷积能够提取更多的特征图信息，类似于网络深度，卷积层越深效果越好，但增加深度会增加模型训练的时间。除此外，相比双层卷积，深度可分离卷积和注意力机制对特征图的语义提取能力强，但对低维细节信息提取差，因此试验中双层卷积模型细节处理效果高于深度可分离卷积和注意力机制。在对比特征融合方式上，特征图相加相比于特征图拼接效果更好，当使用拼接时会在解码器卷积中增加通道维度致使特征提取效果差，而特征图相加是在编码器特征图基础上将提取的多尺度特征图进行加和，能够加强原始细节特征，弥补信息缺失。经过消融实验验证，最佳模型是UNet+Conv_2+Add的改进算法。

4.2 对比试验

本研究对比了PSPNet、SegNet、DeeplabV3plus和改进模型（UNet+Conv_2+Add），每个模型都使用相同配置训练100 epochs后，所有模型损失均达到收敛，并基于测试集测试模型训练效果，获取各

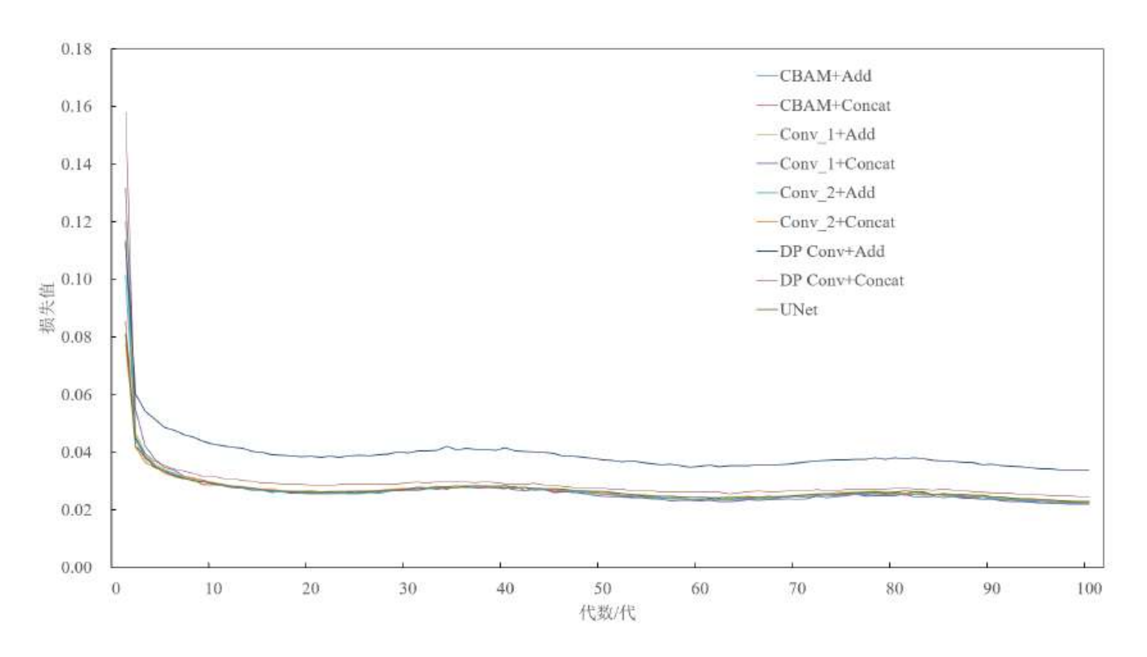


图6 根系图像分割消融实验各种改进模型的损失

Fig. 6 Loss of various improved models in ablation experiments of root image segmentation

项评估指标。模型指标数值对比结果如表4所示。本研究提出的改进模型评估指标最好，mIoU、mRecall和根系F_1值分别为81.62%、86.90%和78.38%。

表4 对比试验中各对比模型评估指标

Table 4 Evaluation indicators of each comparative model in comparative experiments

估计指标	DeeplabV3Plus	PSPNet	SegNet	改进模型(UNet+Conv_2+Add)
Root IoU/%	64.00	54.33	63.08	64.44
Background IoU/%	98.79	98.53	98.79	98.79
mIoU/%	81.39	76.43	89.93	81.62
Root Recall/%	73.53	59.51	73.86	74.25
Background Recall/%	99.47	99.72	99.55	99.55
mRecall/%	86.50	79.61	86.71	86.90
Root Precision/%	81.18	86.17	82.87	83.00
Background Precision/%	99.31	98.81	99.23	99.24
mPrecision/%	90.24	92.49	91.05	91.12
Root F_1/%	77.17	70.40	78.11	78.38
Background F_1/%	99.39	99.26	99.39	99.39

注：R代表根系；B代表背景

对比实际分割图像，改进最佳模型在一些细小的根系分割更加精确，如根系白且短的根（图（7）蓝色框）；在根与根的交界处土壤孔隙的识别也更精确（图（7）红色框）这是因为增加了多尺度特征提取模块使得根系识别更加精细。除此之外，对于颜色深且与土壤颗粒对比度低的根，改进模型仅能实现根的识别且识别稀疏（图（7）黄色框），其余模型识别上下文信息更多，得到的根系更长，改进模型在精确度上的增加也一定程度牺牲了上下文信息提取能力，致使根系在对比度低的情况下，根系分割语义理解力不强。对比图像如图7所示，细节图像如图8所示。

为了进一步评估改进模型和其他卷积模型分割性能，在测试集中随机选择25张图像进行根系表

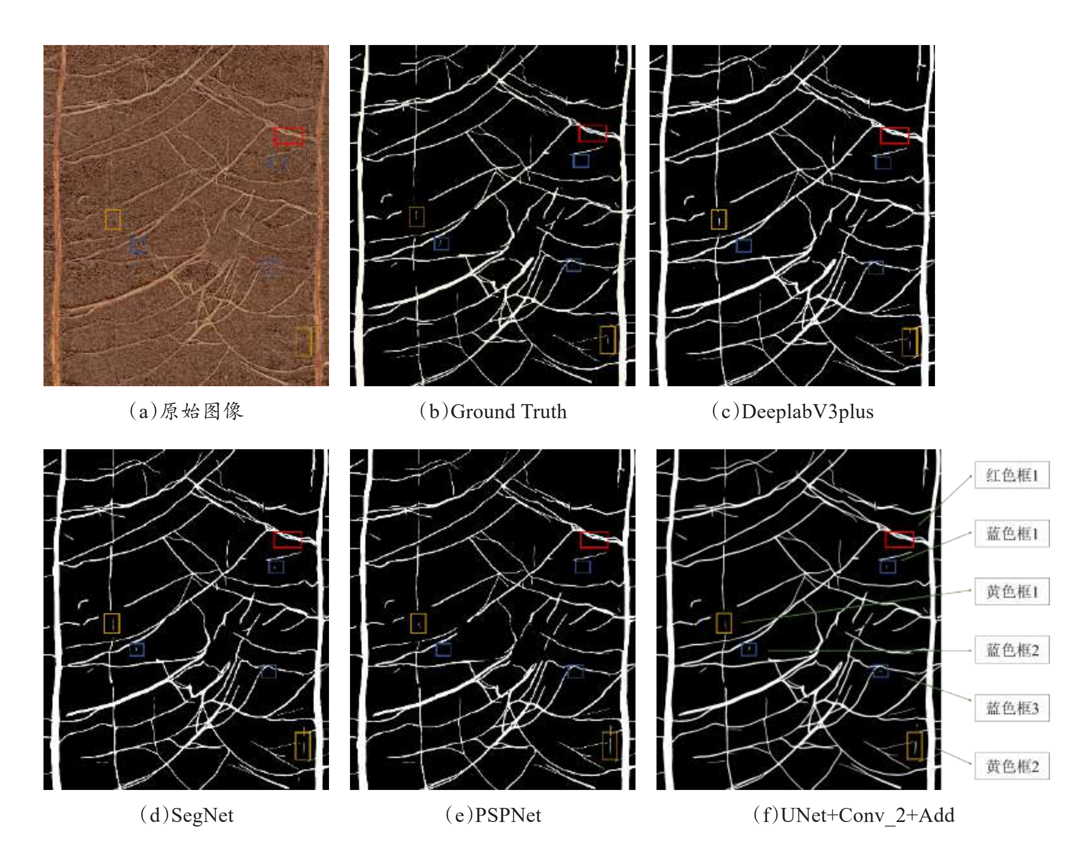


图7 对比试验的各模型分割结果图像

Fig. 7 Images of segmentation results of various models in comparative experiments

型指标测定。测定的四种指标为总根长、平均直径、表面积和容量。测试使用软件为Rhizovision[32]。各模型的定量指标结果如表5所示。在平均直径和表面积两个指标上改进模型与人工标注指标差值最小，总根长和容量两个指标上SegNet的结果最接近人工标注。

综合各项指标，UNet+Conv_2+Add算法效果优势大。对比PSPNet、DeeplabV3Plus算法，模型结构中均使用了额外的多尺度的特征提取方法，SegNet则使用对称的编码解码结构，改进算法则是结合两种结构使得根系分割更精确，但在遮蔽根系识别上一定程度上牺牲信息提取能力。

4.3 迁移学习

单一数据集和混合数据集两者都是关于根系的图像集合，区别在于混合数据包含根系图像种类更加丰富和多样，适用于迁移学习的方法。两种对比模型为UNet原始模型和UNet改进最佳模型（UNet+Conv_2+Add），对比的两种训练方法为普通训练（基于混合数据集，采用数据增强方法，未使用单一数据集训练的权重）和迁移学习（基于混合数据集，采用数据增强方法，使用单一数据集预训练权重）。其训练损失如图9所示，使用迁移学习的初代损失值相比于普通训练更小，这是由于使用了预训练权重的模型在训练时只需要参数微调就可以

表5 根系表型测定的四种表型数据指标

Table 5 Four phenotypic data indicators for root phenotype determination

方法	总根长/px	平均直径/px	容量/px^3	表面积/px^2
手工标注	281,884.9367	16.4984	86,505,316.2980	13,216,115.0220
改进模型UNet+Conv_2+Add	236,648.6779	16.2529	90,592,259.8600	13,275,772.0500
PSPNet	186,125.1123	14.1353	61,695,138.3499	9,377,353.0364
SegNet	240,006.0245	15.7012	85,858,025.2235	12,975,598.3651
DeeplabV3Plus	225,178.9484	15.6688	78,377,863.8983	12,039,045.5469

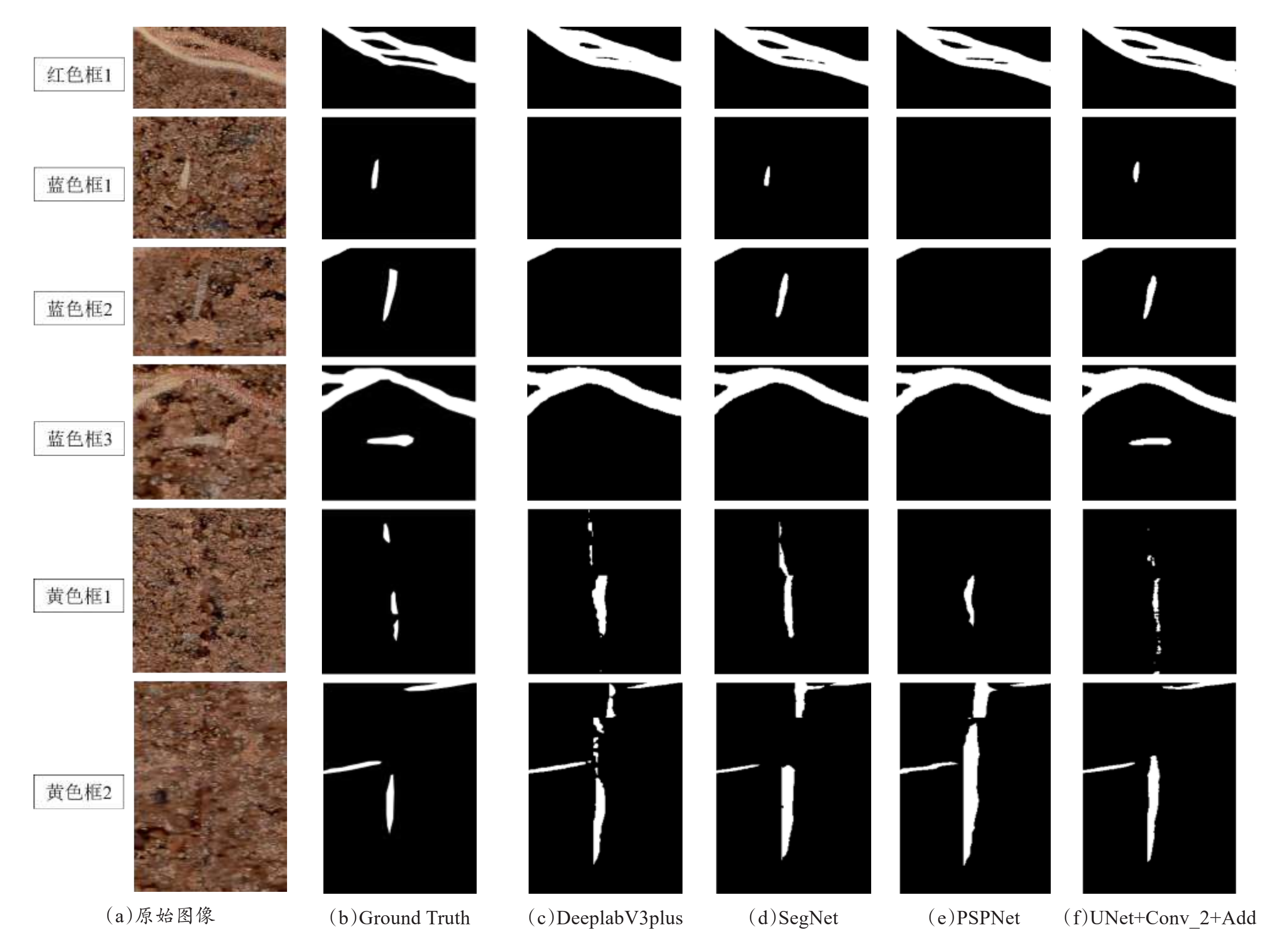


图8　对比试验的各模型对比结果细节小图

Fig. 8 Detailed comparison results of various models in comparative experiments

达到很好的拟合，但普通训练的随机初始化权重使得拟合更加缓慢。除此外，两种训练方式在整个训练周期上的损失值变化也存在差异，迁移学习相比于普通训练收敛更快，且整体损失值更低。

基于测试集得到了两种训练方法的各项评估指标，结果如表6所示。对比结果发现基于迁移学习的两种模型都优于普通训练方法（除Precision外），其中改进模型的迁移学习相比于普通训练的根系IoU提升1.25%、根系Recall提升1.79%、根系F_1值提升0.92%；原始模型的迁移学习相比于普通训练在根系IoU提升0.29%、根系Recall提升0.83%、根系F_1值提升0.21%，结合两者提升发现迁移学习对本研究提出的改进模型提升幅度大、训练效果好。而在迁移学习方式的不同模型评估指标上，本研究提出的改进模型相比原始模型在根系IoU提升1.36%、根系的Recall提升1.99%，根系的F_1提升1.01%；普通训练方式下改进模型相比原始模型在根系IoU提升0.4%、根系的Recall提升1.03%，根系的F_1提升0.30%，训练结果与消融试验结果相拟合，证明迁移学习下的改进模型根系分割更精确。

改进模型（UNet+Conv_2+Add）迁移学习训练后图像分割结果如图10所示。模型能够很好地识别多种土壤不同深度多作物根系，虽然在细小根系处存在误差，但长根的识别相比人工标注更加平滑。

5　结　论

传统算法局限于图像的质量和噪声、土壤环境复杂和算法的泛化能力，深度学习模型能够很好地替代传统方法实现根系的高效、自动提取。因此本研究基于实验室采集的棉花根系数据集和开源多作物微根管数据集两种数据集为试验对象，开展根系

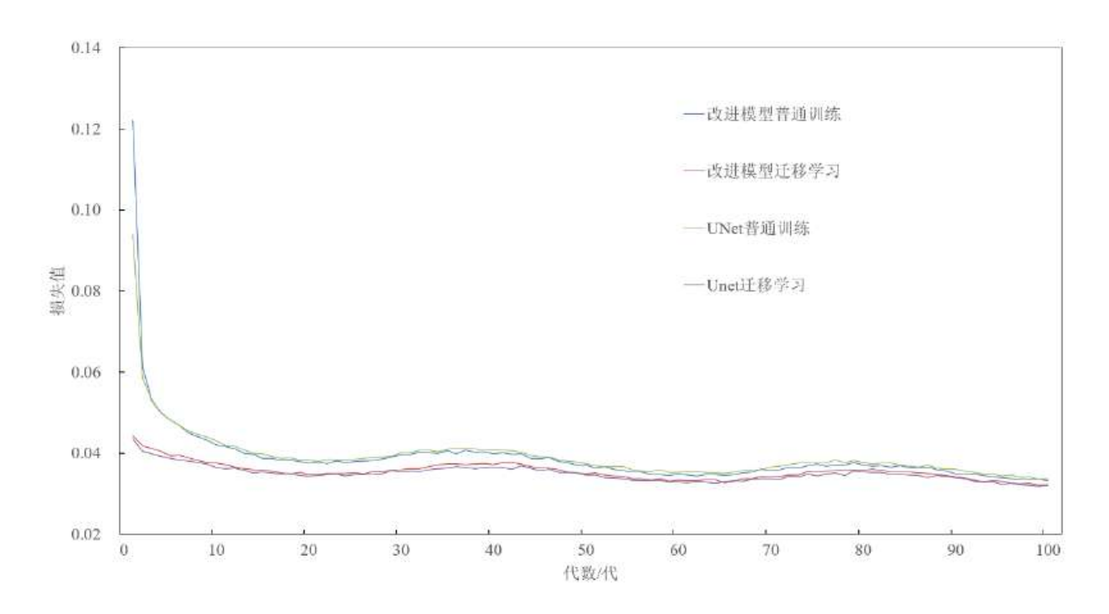


图9 改进模型(UNet+Conv_2+Add)和原始模型在普通训练和迁移学习下的损失

Fig. 9 Comparison of the loss of improved model(UNet+Conv_2+Add) and the original model under ordinary training and transfer learning

表6 改进模型(UNet+Conv_2+Add)和原始模型在普通训练和迁移学习下各项评估指标

Table 6 Evaluation indicators of improved model(UNet+Conv_2+Add) and original model under ordinary training and transfer learning

评估指标	UNet普通训练	UNet迁移学习	改进模型普通训练	改进模型迁移学习
R IoU/%	62.93	63.22	63.33	64.58
B IoU/%	98.75	98.75	98.76	98.79
mIoU/%	80.84	80.99	81.40	81.68
R Recall/%	71.27	72.10	72.30	74.09
B Recall/%	99.60	99.58	99.58	99.56
mRecall/%	85.44	85.84	85.94	86.83
R Precision/%	84.32	83.69	83.62	83.41
B Precision/%	99.14	99.17	99.17	99.23
mPrecision/%	91.73	91.43	91.40	91.32
R F_1/%	77.25	77.46	77.55	78.47
B F_1/%	99.37	99.37	99.37	99.39

注：R代表根系；B代表背景

分割工作，主要研究结果如下。

（1）提出一种基于UNet的改进模型，加入了多尺度图像提取融合模块实现不同尺寸图像细节特征提取，提高模型识别根系的能力。并通过消融实验验证该模块的有效性。改进后的模型（UNet+Conv_2+Add）相比其他模型和其他改进算法得到的综合评价指标较好，mIoU、mRecall和根系F_1值分别为81.62%、86.90%和78.39%，相比于原始UNet的mIoU、mRecall和根系F_1值提升0.37%、0.99%和0.56%，证明改进模型具有优势性。

（2）试验时采用数据增强方式（裁剪、旋转、翻转、改变图像的属性四种增强方法）实现数据的多样性扩充，基于两种数据集完成迁移学习。通过改进后的损失值和各项评估指标对比验证迁移学习的优越性，根系IoU提升1.25%、根系Recall提升1.79%、根系F_1值提升0.92%。结果证明采用迁移学习的方法收敛速度快，能够进一步增加模型的分割性能。

研究结果证明改进后模型不仅优于原始模型，并且在采用迁移学习后增加模型泛化能力。但是模型还存在不足之处。一是模型的参数巨大，对于模型部署存在限制，后续应该考虑简化模型。二是样本类型种类数量问题，后续会收集其他作物的根系图像进一步验证模型有效性。

利益冲突声明：本研究不存在研究者以及与公开研究成果有关的利益冲突。

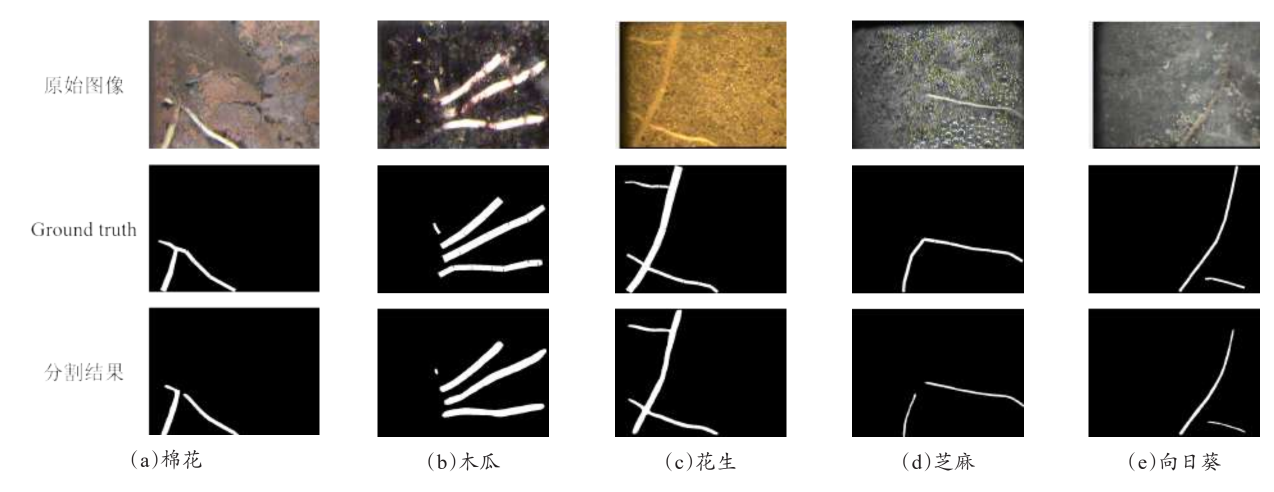


图10 改进模型(UNet+Conv_2+Add)的迁移学习分割结果

Fig. 10 Transfer learning segmentation results for improved models (UNet+Conv_2+Add)

参考文献：

[1] 王宁, 李继光, 娄翼来, 等. 作物根系形态对施肥措施的响应[J]. 中国农学通报, 2020, 36(3): 53-58.
WANG N, LI J G, LOU Y L, et al. Response of crop root morphology to fertilization measures[J]. Chinese agricultural science bulletin, 2020, 36(3): 53-58.

[2] DONG H Z, NIU Y H, LI W J, et al. Effects of cotton rootstock on endogenous cytokinins and abscisic acid in xylem sap and leaves in relation to leaf senescence[J]. Journal of experimental botany, 2008, 59(6): 1295-1304.

[3] 吴茜, 张伟欣, 张玲玲, 等. 植物根系表型信息获取技术研究进展[J]. 江苏农业科学, 2021, 49(5): 31-37.
WU Q, ZHANG W X, ZHANG L L, et al. Research progress on acquisition of plant root phenotype information[J]. Jiangsu agricultural sciences, 2021, 49(5): 31-37.

[4] ZHANG B W. Plant root research methods and trends[J]. Agricultural science & technology, 2017, 18(12): 2295-2298, 2302.

[5] 肖爽, 刘连涛, 张永江, 等. 植物微根系原位观测方法研究进展[J]. 植物营养与肥料学报, 2020, 26(2): 370-385.
XIAO S, LIU L T, ZHANG Y J, et al. Review on new methods of *in situ* observation of plant micro-roots and interpretation of root images[J]. Journal of plant nutrition and fertilizers, 2020, 26(2): 370-385.

[6] 赵先丽, 蔡福, 李荣平, 等. 春玉米根系图像语义分割最佳分辨率和概率阈值研究[J]. 核农学报, 2023, 37(8): 1690-1699.
ZHAO X L, CAI F, LI R P, et al. Optimal resolution and probability threshold for the semantic segmentation of spring maize root image[J]. Journal of nuclear agricultural sciences, 2023, 37(8): 1690-1699.

[7] 何勇, 李禧尧, 杨国峰, 等. 室内高通量种质资源表型平台研究进展与展望[J]. 农业工程学报, 2022, 38(17): 127-141.
HE Y, LI X Y, YANG G F, et al. Research progress and prospect of indoor high-throughput germplasm phenotyping platforms[J]. Transactions of the Chinese society of agricultural engineering, 2022, 38(17): 127-141.

[8] PERELMAN A, LAZAROVITCH N, VANDERBORGHT J, et al. Quantitative imaging of sodium concentrations in soil-root systems using magnetic resonance imaging (MRI)[J]. Plant and soil, 2020, 454(1/2): 171-185.

[9] SCOTSON C, DUNCAN S, WILLIAMS K, et al. X-ray computed tomography imaging of solute movement through ridged and flat plant systems[J]. European journal of soil science, 2021, 72 (1): 198-214

[10] HAMMAC W A, PAN W L, BOLTON R P, et al. High resolution imaging to assess oilseed species' root hair responses to soil water stress[J]. Plant and soil, 2011, 339(1/2): 125-135.

[11] MOHAMED A, MONNIER Y, MAO Z, et al. An evaluation of inexpensive methods for root image acquisition when using rhizotrons[J]. Plant methods, 2017, 13(1): 1-13.

[12] ZHAO H J, WANG N, SUN H C, et al. RhizoPot platform: A high-throughput *in situ* root phenotyping platform with integrated hardware and software[J]. Frontiers in plant science, 2022, 13: ID 1004904.

[13] DAS A, SCHNEIDER H, BURRIDGE J, et al. Digital imaging of root traits (DIRT): A high-throughput computing and collaboration platform for field-based root phenomics[J]. Plant methods, 2015, 11: ID 51.

[14] GALKOVSKYI T, MILEYKO Y, BUCKSCH A, et al. GiA Roots: Software for the high throughput analysis of plant root system architecture[J]. BMC plant biology, 2012, 12: ID 116.

[15] PIERRET A, GONKHAMDEE S, JOURDAN C, et al. IJ_Rhizo: An open-source software to measure scanned images of root samples[J]. Plant and soil, 2013, 373(1/2): 531-539.

[16] ARMENGAUD P, ZAMBAUX K, HILLS A, et al. EZ-Rhizo: Integrated software for the fast and accurate measurement of root system architecture[J]. The plant journal, 2009, 57(5): 945-956.

[17] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[EB/OL]. arX-

iv: 1411.4038, 2014.

[18] KAMAL S, SHENDE V G, SWAROOPA K, et al. FCN network-based weed and crop segmentation for IoT-aided agriculture applications[J]. Wireless communications and mobile computing, 2022, 2022: 1-10.

[19] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation[EB/OL]. arXiv: 1511.00561, 2015.

[20] WANG T, ROSTAMZA M, SONG Z H, et al. SegRoot: A high throughput segmentation method for root image analysis[J]. Computers and electronics in agriculture, 2019, 162: 845-854.

[21] ZHAO H S, SHI J P, QI X J, et al. Pyramid scene parsing network[EB/OL]. arXiv: 1612.01105, 2016.

[22] ZHANG R, CHEN J, FENG L, et al. A Refined Pyramid Scene Parsing Network for Polarimetric SAR Image Semantic Segmentation in Agricultural Areas[J]. IEEE Geoscience and Remote Sensing Letters, 2022, 19: 1-5.

[23] CHEN L C, ZHU Y K, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[EB/OL]. arXiv: 1802.02611. 2018.

[24] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation[EB/OL]. arXiv: 1706.05587, 2017.

[25] KANG J, LIU L T, ZHANG F C, et al. Semantic segmentation model of cotton roots *in situ* image based on attention mechanism[J]. Computers and electronics in agriculture, 2021, 189: ID 106370.

[26] SHEN C, LIU L T, ZHU L X, et al. High-throughput *in situ* root image segmentation based on the improved DeepLabv3+ method[J]. Frontiers in plant science, 2020, 11: ID 576791.

[27] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[EB/OL]. arXiv: 1505.04597, 2015.

[28] 林娜, 何静, 王斌, 等. 结合植被光谱特征与Sep-UNet的城市植被信息智能提取方法[J]. 地球信息科学学报, 2023, 25(8): 1717-1729.
LIN N, HE J, WANG B, et al. Intelligent extraction of urban vegetation information based on vegetation spectral signature and sep-UNet[J]. Journal of geo-information science, 2023, 25(8): 1717-1729.

[29] 申传庆, 王凯, 王文杰. 基于ResNet-UNet的地表覆盖自动分类技术研究[J]. 地理空间信息, 2023, 21(6): 21-23, 27.
SHEN C Q, WANG K, WANG W J. Research on automatic classification technology of land coverage based on ResNet-UNet[J]. Geospatial information, 2023, 21(6): 21-23, 27.

[30] 陈桂芬, 赵姗, 曹丽英, 等. 基于迁移学习与卷积神经网络的玉米植株病害识别[J]. 智慧农业, 2019, 1(2): 34-44.
CHEN G F, ZHAO S, CAO L Y, et al. Corn plant disease recognition based on migration learning and convolutional neural network[J]. Smart agriculture, 2019, 1(2): 34-44.

[31] XU W H, YU G H, CUI Y M, et al. PRMI: A dataset of minirhizotron images for diverse plant root study[EB/OL]. arXiv: 2201.08002, 2022.

[32] SEETHEPALLI A, DHAKAL K, GRIFFITHS M, et al. RhizoVision Explorer: Open-source software for root image analysis and measurement standardization[J]. AoB PLANTS, 2021, 13(6): ID plab056.

Root Image Segmentation Method Based on Improved UNet and Transfer Learning

TANG Hui[1], WANG Ming[2], YU Qiushi[1], ZHANG Jiaxi[1], LIU Liantao[3], WANG Nan[1*]

(1. *College of mechanical and electrical engineering*, *Hebei Agricultural University*, *Baoding* 071001, *China*;
2. *Hebei Education Examinations Authority*, *Shijiazhuang* 050091, *China*;
3. *College of agronomy*, *Hebei Agricultural University*, *Baoding* 071001, *China*)

Abstract:

[Objective] The root system is an important component of plant composition, and its growth and development are crucial for plants. Root image segmentation is an important method for obtaining root phenotype information and analyzing root growth patterns. Research on root image segmentation still faces difficulties, because of the noise and image quality limitations, the intricate and diverse soil environment, and the ineffectiveness of conventional techniques. This paper proposed a multi-scale feature extraction root segmentation algorithm that combined data augmentation and transfer learning to enhance the generalization and universality of the root image segmentation models in order to increase the speed, accuracy, and resilience of root image segmentation.

[Methods] Firstly, the experimental datasets were divided into a single dataset and a mixed dataset. The single dataset acquisition was obtained from the experimental station of Hebei Agricultural University in Baoding city. Additionally, a self-made RhizoPot device

was used to collect images with a resolution pixels of 10,200×14,039, resulting in a total of 600 images. In this experiment, 100 sheets were randomly selected to be manually labeled using Adobe Photoshop CC2020 and segmented into resolution pixels of 768×768, and divided into training, validation, and test sets according to 7:2:1. To increase the number of experimental samples, an open source multi-crop mixed dataset was obtained in the network as a supplement, and it was reclassified into training, validation, and testing sets. The model was trained using the data augmentation strategy, which involved performing data augmentation operations at a set probability of 0.3 during the image reading phase, and each method did not affect the other. When the probability was less than 0.3, changes would be made to the image. Specific data augmentation methods included changing image attributes, randomly cropping, rotating, and flipping those images. The UNet structure was improved by designing eight different multi-scale image feature extraction modules. The module structure mainly included two aspects: Image convolution and feature fusion. The convolution improvement included convolutional block attention module (CBAM), depthwise separable convolution (DP Conv), and convolution (Conv). In terms of feature fusion methods, improvements could be divided into concatenation and addition. Subsequently, ablation tests were conducted based on a single dataset, data augmentation, and random loading of model weights, and the optimal multi-scale feature extraction module was selected and compared with the original UNet. Similarly, a single dataset, data augmentation, and random loading of model weights were used to compare and validate the advantages of the improved model with the PSPNet, SegNet, and DeeplabV3Plus algorithms. The improved model used pre-trained weights from a single dataset to load and train the model based on mixed datasets and data augmentation, further improving the model's generalization ability and root segmentation ability.

[Results and Discussions] The results of the ablation tests indicated that Conv_ 2+Add was the best improved algorithm. Compared to the original UNet, the mIoU, mRecall, and root F_1 values of the model increased by 0.37%, 0.99%, and 0.56%, respectively. And, comparative experiments indicate Unet+Conv_2+Add model was superior to the PSPNet, SegNet, and DeeplabV3Plus models, with the best evaluation results. And the values of mIoU, mRecall, and the harmonic average of root F_1 were 81.62%, 86.90%, and 77.97%, respectively. The actual segmented images obtained by the improved model were more finely processed at the root boundary compared to other models. However, for roots with deep color and low contrast with soil particles, the improved model could only achieve root recognition and the recognition was sparse, sacrificing a certain amount of information extraction ability. This study used the root phenotype evaluation software Rhizovision to analyze the root images of the Unet+Conv_2+Add improved model, PSPNet, SegNet, and DeeplabV3Plu, respectively, to obtain the values of the four root phenotypes (total root length, average diameter, surface area, and capacity), and the results showed that the average diameter and surface area indicator values of the improved model, Unet+Conv_2+Add had the smallest differences from the manually labeled indicator values and the SegNet indicator values for the two indicators. Total root length and volume were the closest to those of the manual labeling. The results of transfer learning experiments proved that compared with ordinary training, the transfer training of the improved model UNet+Conv_2+Add increased the IoU value of the root system by 1.25%. The Recall value of the root system was increased by 1.79%, and the harmonic average value of F_1 was increased by 0.92%. Moreover, the overall convergence speed of the model was fast. Compared with regular training, the transfer training of the original UNet improved the root IoU by 0.29%, the root Recall by 0.83%, and the root F_1 value by 0.21%, which indirectly confirmed the effectiveness of transfer learning.

[Conclusions] The multi-scale feature extraction strategy proposed in this study can accurately and efficiently segment roots, and further improve the model's generalization ability using transfer learning methods, providing an important research foundation for crop root phenotype research.

Key words: deep learning; root image segmentation; UNet; multi-scale characteristics; transfer learning

Foundation Items: Top-notch Talent Plan Program of the Education Department of Hebei Province (BJ2021058); Central Guiding Local Science and Technology Development Fund Projects(236Z7402G); State Key Laboratory of North China Crop Improvement and Regulation(NCCIR2021ZZ-23)